

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, C07K 14/47, 16/18, C12Q 1/68, C12N 15/11, A61K 38/17, G01N 33/68	A1	(11) International Publication Number: WO 98/42839 (43) International Publication Date: 1 October 1998 (01.10.98)
(21) International Application Number: PCT/US98/05871 (22) International Filing Date: 25 March 1998 (25.03.98) (30) Priority Data: 08/824,873 26 March 1997 (26.03.97) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 08/824,873 (CIP) Filed on 26 March 1997 (26.03.97) (71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). (74) Agent: BILLINGS, Lucy, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).		(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, GB, IL, JP, KR, MX, NO, NZ, RU, SE, SG, US, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: HUMAN RAB PROTEIN, SRAB (57) Abstract The present invention provides a human RAB protein (SRAB) and polynucleotides which identify and encode SRAB. The invention also provides expression vectors, host cells, agonists, antibodies, and antagonists. The invention also provides methods for treating disorders associated with expression of SRAB.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

HUMAN RAB PROTEIN, SRAB

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of novel RAB protein and to the use of these sequences in the diagnosis, prevention, and treatment of disorders associated with vesicle trafficking, viral infection, and cancer.

BACKGROUND ART

Vesicle trafficking is defined as the vesicular transport of materials between different subcellular compartments of eukaryotic cells. Vesicles bud from a donor membrane and fuse with a recipient one carrying internalized materials from one site to another. Rab proteins, low molecular weight guanine triphosphatases (GTPases) of the Ras superfamily, help regulate vesicular transport by directing the vesicles to and from the correct membrane surfaces (Novick, P. and Brennwald, P. (1993) Cell 75:597-601).

In their role as GTP transferases, Rab proteins are prenylated, bind GTP and are placed in particular intracellular membranes by escort proteins (Khosravi-Far, R. et al. (1991) Proc. Natl. Acad. Sci. 88:6264-6268). Experimental evidence shows that in the vesicle, GTP-bound Rab proteins interact with SNARE factors to direct vesicle transport. After transport, GTPase activating proteins in the target membrane convert Rab proteins to the GDP-bound form, and guanine-nucleotide dissociation inhibitor helps return the GDP-bound protein to its membrane of origin.

Evidence indicates that pancreatic acinar cell enzyme secretion may be regulated by Rab proteins. Wagner, A. et al. (1995, Biochem. Biophys. Res. Comm. 27:950-956) used a Rab 3A probe to screen a rat pancreatic cDNA library and identified Rab 26, a new 190 amino acid member of the Rab protein family. Rab 26 contains conserved GTP-binding regions and C-terminal isoprenylation sites characteristic of the Rab protein family members. Northern analysis of Rab 26 demonstrated expression in kidney, brain, salivary gland, lung, and pancreas. The authors speculate that Rab 26 might be involved in the vesicle trafficking which regulates pancreatic secretion

To date, more than 30 Rab proteins have been identified, and each may have a characteristic intracellular location where it functions in distinct, tissue-specific transport events. Some examples of Rab proteins, their functions, and implication in disease processes are summarized. Rab3A is a component of brain synaptic vesicles and has been implicated in the regulation of neurotransmitter release (Matteoli, M. et al. (1991) J. Cell Biol. 115:625-33). Overexpression of Rab proteins significantly enhances the function of Rev, a viral gene essential

for processing HIV-1 (Fridell, R.A. et al. (1996) Proc. Natl. Acad. Sci. 93:4421-4424). A deficiency in the prenylation of one particular Rab is associated with choroideremia, a form of retinal degeneration that causes blindness (Seabra, M. et al. (1993) Science 259:377-81).

Interaction between Rab protein and Cdc2 protein kinase in vitro inhibited vesicle fusion and
5 implicated Rab protein function in mediating cell cycle events (Toumikoski, T. et al. (1989) Nature 342:942-945). Thus, Rab proteins appear to be involved in the complex and critical processes of vesicle trafficking for the directed release of various molecules including digestive enzymes and signaling molecules, for viral processing, and in mitotic and meiotic cell cycle.

The discovery of polynucleotides encoding novel Rab proteins and the molecules
10 themselves satisfy a need in the art by providing new compositions useful in the diagnosis, prevention, or treatment of disorders associated with vesicle trafficking, viral infection, and cancer.

DISCLOSURE OF THE INVENTION

The present invention features a novel Rab protein hereinafter designated SRAB and
15 characterized as having similarity to the rat Rab 26 (GI 619734) and the human Rab 8 protein (GI 234746).

Accordingly, the invention features a substantially purified SRAB having the amino acid sequence shown in SEQ ID NO:1.

One aspect of the invention features isolated and substantially purified polynucleotides
20 that encode SRAB. In a particular aspect, the polynucleotide is the nucleotide sequence of SEQ ID NO:2.

The invention also relates to a polynucleotide sequence comprising the complement of SEQ ID NO:2 or variants thereof. In addition, the invention features polynucleotide sequences which hybridize under stringent conditions to SEQ ID NO:2.

25 The invention additionally features nucleic acid sequences encoding polypeptides, oligonucleotides, peptide nucleic acids (PNA), fragments, portions or antisense molecules thereof, and expression vectors and host cells comprising polynucleotides that encode SRAB. The present invention also features antibodies which bind specifically to SRAB, and pharmaceutical compositions comprising substantially purified SRAB. The invention also
30 features agonists and antagonists of SRAB, and methods for diagnosing, preventing or treating SRAB associated disorders using the nucleotide, the protein and antagonists presented herein.

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A, 1B, 1C and 1D show the amino acid sequence (SEQ ID NO:1) and nucleic

acid sequence (SEQ ID NO:2) of SRAB. The alignment was produced using MacDNASIS PRO™ software (Hitachi Software Engineering Co., Ltd., San Bruno, CA).

Figures 2A and 2B shows the amino acid sequence alignments among SRAB (SEQ ID NO:1), rat Rab 26 (GI 619734; SEQ ID NO:3), and human Rab 8 (GI 234746; SEQ ID NO:4).

5 The alignment was produced using the multisequence alignment program of DNASTAR™ software (DNASTAR Inc, Madison WI).

MODES FOR CARRYING OUT THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines,
10 vectors, and reagents described as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for
15 example, reference to "a host cell" includes a plurality of such host cells, reference to the "antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention
20 belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are reported in the publications which might be used in connection with the invention.
25 Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic
30 origin which may be single- or double-stranded, and represent the sense or antisense strand. Similarly, "amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules.

Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms, such as "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

5 "Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P.E. et al. (1993) Anticancer Drug Des. 8:53-63).

SRAB, as used herein, refers to the amino acid sequences of substantially purified SRAB
10 obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source whether natural, synthetic, semi-synthetic, or recombinant.

"Consensus", as used herein, refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, or which has been extended using XL-PCR™ (Perkin
15 Elmer, Norwalk, CT) in the 5' and/or the 3' direction and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte clone using the GELVIEW™ Fragment Assembly system (GCG, Madison, WI), or which has been both extended and assembled.

A "variant" of SRAB, as used herein, refers to an amino acid sequence that is altered by
20 one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted,
25 inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNASTAR software.

A "deletion", as used herein, refers to a change in either amino acid or nucleotide sequence in which one or more amino acid or nucleotide residues, respectively, are absent.

An "insertion" or "addition", as used herein, refers to a change in an amino acid or
30 nucleotide sequence resulting in the addition of one or more amino acid or nucleotide residues, respectively, as compared to the naturally occurring molecule.

A "substitution", as used herein, refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

The term "biologically active", as used herein, refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic SRAB, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells
5 and to bind with specific antibodies.

The term "agonist", as used herein, refers to a molecule which, when bound to SRAB, causes a change in SRAB which modulates the activity of SRAB. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to SRAB.

The terms "antagonist" or "inhibitor", as used herein, refer to a molecule which, when
10 bound to SRAB, blocks or modulates the biological or immunological activity of SRAB. Antagonists and inhibitors may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to SRAB.

The term "modulate", as used herein, refers to a change or an alteration in the biological activity of SRAB. Modulation may be an increase or a decrease in protein activity, a change in
15 binding characteristics, or any other change in the biological, functional or immunological properties of SRAB.

The term "mimetic", as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of SRAB or portions thereof and, as such, is able to effect some or all of the actions of Rab protein.

20 The term "derivative", as used herein, refers to the chemical modification of a nucleic acid encoding SRAB or the encoded SRAB. Illustrative of such modifications would be replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative would encode a polypeptide which retains essential biological characteristics of the natural molecule.

The term "substantially purified", as used herein, refers to nucleic or amino acid
25 sequences that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

"Amplification" as used herein refers to the production of additional copies of a nucleic acid sequence and is generally carried out using polymerase chain reaction (PCR) technologies
30 well known in the art (Dieffenbach, C.W. and G.S. Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, NY).

The term "hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex", as used herein, refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen binds between complementary G and C bases and between complementary A and T bases; these hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences
5 hydrogen bond in an antiparallel configuration. A hybridization complex may be formed in solution (e.g., C_{0t} or R_{0t} analysis) or between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., membranes, filters, chips, pins or glass slides to which cells have been fixed for in situ hybridization).

The terms "complementary" or "complementarity", as used herein, refer to the natural
10 binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, for the sequence "A-G-T" binds to the complementary sequence "T-C-A". Complementarity between two single-stranded molecules may be "partial", in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has
15 significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands.

The term "similarity", as used herein, refers to a degree of complementarity. There may be partial similarity or complete similarity (i.e., identity). A partially complementary sequence is
20 one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid; it is referred to using the functional term "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially similar sequence or probe will compete for and
25 inhibit the binding (i.e., the hybridization) of a completely similar sequence or probe to the target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial
30 degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence.

As known in the art, numerous equivalent conditions may be employed to comprise either low or high stringency conditions. Factors such as the length and nature (DNA, RNA, base

composition) of the sequence, nature of the target (DNA, RNA, base composition, presence in solution or immobilization, etc.), and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulfate and/or polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high
5 stringency different from, but equivalent to, the above listed conditions.

The term "stringent conditions", as used herein, is the "stringency" which occurs within a range from about $T_m - 5^\circ\text{C}$ (5°C below the melting temperature (T_m) of the probe) to about 20°C to 25°C below T_m . As will be understood by those of skill in the art, the stringency of hybridization may be altered in order to identify or detect identical or related polynucleotide
10 sequences.

The term "antisense", as used herein, refers to nucleotide sequences which are complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest
15 in a reverse orientation to a viral promoter which permits the synthesis of a complementary strand. Once introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further transcription or translation. In this manner, mutant phenotypes may be generated. The designation "negative" is sometimes used in reference to the antisense strand, and "positive" is sometimes used in
20 reference to the sense strand.

The term "portion", as used herein, with regard to a protein (as in "a portion of a given protein") refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid. Thus, a protein "comprising at least a portion of the amino acid sequence of SEQ ID NO:1" encompasses the
25 full-length human SRAB and fragments thereof.

"Transformation", as defined herein, describes a process by which exogenous DNA enters and changes a recipient cell. It may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is
30 selected based on the host cell being transformed and may include, but is not limited to, viral infection, electroporation, lipofection, and particle bombardment. Such "transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. They also include cells

which transiently express the inserted DNA or RNA for limited periods of time.

The term "antigenic determinant", as used herein, refers to that portion of a molecule that makes contact with a particular antibody (i.e., an epitope). When a protein or fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the
5 production of antibodies which bind specifically to a given region or three-dimensional structure on the protein; these regions or structures are referred to as antigenic determinants. An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The terms "specific binding" or "specifically binding", as used herein, in reference to the
10 interaction of an antibody and a protein or peptide, mean that the interaction is dependent upon the presence of a particular structure (i.e., the antigenic determinant or epitope) on the protein; in other words, the antibody is recognizing and binding to a specific protein structure rather than to proteins in general. For example, if an antibody is specific for epitope "A", the presence of a protein containing epitope A (or free, unlabeled A) in a reaction containing labeled "A" and the
15 antibody will reduce the amount of labeled A bound to the antibody.

The term "sample", as used herein, is used in its broadest sense. A biological sample suspected of containing nucleic acid encoding SRAB or fragments thereof may comprise a cell, chromosomes isolated from a cell (e.g., a spread of metaphase chromosomes), genomic DNA (in solution or bound to a solid support such as for Southern analysis), RNA (in solution or bound to
20 a solid support such as for northern analysis), cDNA (in solution or bound to a solid support), an extract from cells or a tissue, and the like.

The term "correlates with expression of a polynucleotide", as used herein, indicates that the detection of the presence of ribonucleic acid that is similar to SEQ ID NO:2 by northern analysis is indicative of the presence of mRNA encoding SRAB in a sample and thereby
25 correlates with expression of the transcript from the polynucleotide encoding the protein.

"Alterations" in the polynucleotide of SEQ ID NO:2, as used herein, comprise any alteration in the sequence of polynucleotides encoding SRAB including deletions, insertions, and point mutations that may be detected using hybridization assays. Included within this definition is the detection of alterations to the genomic DNA sequence which encodes SRAB (e.g., by
30 alterations in the pattern of restriction fragment length polymorphisms capable of hybridizing to SEQ ID NO:2), the inability of a selected fragment of SEQ ID NO:2 to hybridize to a sample of genomic DNA (e.g., using allele-specific oligonucleotide probes), and improper or unexpected hybridization, such as hybridization to a locus other than the normal chromosomal locus for the

polynucleotide sequence encoding SRAB (e.g., using fluorescent in situ hybridization [FISH] to metaphase chromosomes spreads).

As used herein, the term "antibody" refers to intact molecules as well as fragments thereof, such as Fa, F(ab')₂, and Fv, which are capable of binding the epitopic determinant.

5 Antibodies that bind SRAB polypeptides can be prepared using intact polypeptides or fragments containing small peptides of interest as the immunizing antigen. The polypeptide or peptide used to immunize an animal can be derived from the transition of RNA or synthesized chemically, and can be conjugated to a carrier protein, if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin and thyroglobulin. The coupled peptide is
10 then used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

The term "humanized antibody", as used herein, refers to antibody molecules in which amino acids have been replaced in the non-antigen binding regions in order to more closely resemble a human antibody, while still retaining the original binding ability.

THE INVENTION

15 The invention is based on the discovery of a novel Rab protein, SRAB, the polynucleotides encoding SRAB, and the use of these compositions for the diagnosis, prevention and treatment of disorders associated with vesicle trafficking, viral infection, and cancer.

Nucleic acids encoding the human SRAB of the present invention were first identified in Incyte Clone 738957 from the pancreatic cDNA library (PANCNOT04) through a
20 computer-generated search for amino acid sequence alignments. SEQ ID NO:2 was derived from extension of Incyte Clone 738957 (PANCNOT04).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1, as shown in Figures 1A, 1B, 1C and 1D. SRAB is 190 amino acids in length and has 4 potential GTP-binding sites (residues 7-10, 52-58, 111-114, and 139-143).
25 SRAB has a consensus sequence for isoprenylation--CCX in the C-terminus; multiple protein kinase C phosphorylation sites at residues 81, 141, and 163; myristylation sites at residues 7, 15, and 184; and a potential glycosaminoglycan attachment site at residue 6. As shown in Figures 2A and 2B, SRAB has chemical and structural similarity with rat Rab 26 (SEQ ID NO:3) and human Rab 8 (SEQ ID NO:4). In particular, SRAB and rat Rab 26 protein share 91% identity, while
30 SRAB has a 47% identity with human Rab 8. SRAB and rat Rab 26 have rather similar isoelectric points of 8.22 and 8.40, respectively. Northern analysis shows the expression of this sequence in both adult and infant pancreatic cDNA libraries.

The invention also encompasses SRAB variants. A preferred SRAB variant is one having

at least 80%, and more preferably 90%, amino acid sequence identity to the SRAB amino acid sequence (SEQ ID NO:1). A most preferred SRAB variant is one having at least 95% amino acid sequence identity to SEQ ID NO:1.

The invention also encompasses polynucleotides which encode SRAB. Accordingly, any
5 nucleic acid sequence which encodes the amino acid sequence of SRAB can be used to generate recombinant molecules which express SRAB. In a particular embodiment, the invention encompasses the polynucleotide comprising the nucleic acid sequence of SEQ ID NO:2 as shown in Figures 1A, 1B, 1C and 1D.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the
10 genetic code, a multitude of nucleotide sequences encoding SRAB, some bearing minimal similarity to the nucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the
15 nucleotide sequence of naturally occurring SRAB, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode SRAB and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring SRAB under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide
20 sequences encoding SRAB or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SRAB and its derivatives without altering the encoded amino acid sequences include
25 the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences, or portions thereof, which encode SRAB and its derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell
30 systems using reagents that are well known in the art at the time of the filing of this application. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding SRAB or any portion thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of

hybridizing to the claimed nucleotide sequences, and in particular, those shown in SEQ ID NO:2, under various conditions of stringency. Hybridization conditions are based on the melting temperature (T_m) of the nucleic acid binding complex or probe, as taught in Wahl, G.M. and S.L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A.R. (1987; Methods Enzymol. 5 152:507-511), and may be used at a defined stringency.

Altered nucleic acid sequences encoding SRAB which are encompassed by the invention include deletions, insertions, or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent SRAB. The encoded protein may also contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and 10 result in a functionally equivalent SRAB. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of SRAB is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head 15 groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; phenylalanine and tyrosine.

Also included within the scope of the present invention are alleles of the genes encoding SRAB. As used herein, an "allele" or "allelic sequence" is an alternative form of the gene which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered 20 mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

25 Methods for DNA sequencing which are well known and generally available in the art may be used to practice any embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase® (US Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, IL), or combinations of recombinant polymerases and proofreading exonucleases such 30 as the ELONGASE Amplification System marketed by Gibco BRL (Gaithersburg, MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) and the ABI 377 DNA sequencers (Perkin Elmer).

The nucleic acid sequences encoding SRAB may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, "restriction-site" PCR, uses universal primers to retrieve unknown sequence adjacent
5 to a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). In particular, genomic DNA is first amplified in the presence of primer to linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse
10 transcriptase.

Inverse PCR may also be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). The primers may be designed using OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of
15 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA
20 (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may also be used to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before performing PCR.

Another method which may be used to retrieve unknown sequences is that of Parker, J.D.
25 et al. (1991: Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PromoterFinder™ libraries to walk in genomic DNA (Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they
30 will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into the 5' and 3' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection
5 of the emitted wavelengths by a charge coupled device camera. Output/light intensity may be converted to electrical signal using appropriate software (e.g. Genotyper™ and Sequence Navigator™, Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts
10 in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode SRAB, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of SRAB in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the
15 same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express SRAB.

As will be understood by those of skill in the art, it may be advantageous to produce SRAB-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate
20 of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter SRAB encoding sequences for a variety of reasons,
25 including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

30 In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding SRAB may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of SRAB activity, it may be useful to encode a chimeric SRAB protein that can be recognized by a commercially available antibody. A

fusion protein may also be engineered to contain a cleavage site located between the SRAB encoding sequence and the heterologous protein sequence, so that SRAB may be cleaved and purified away from the heterologous moiety.

In another embodiment, sequences encoding SRAB may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of SRAB, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J.Y. et al. (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer).

The newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, NY). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; Creighton, supra). Additionally, the amino acid sequence of SRAB, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a biologically active SRAB, the nucleotide sequences encoding SRAB or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence.

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding SRAB and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, NY, and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding SRAB. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus

expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript® phagemid (Stratagene, LaJolla, CA) or pSport1™ plasmid (Gibco BRL) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO; and storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding SRAB, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for SRAB. For example, when large quantities of SRAB are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional E. coli cloning and expression vectors such as Bluescript® (Stratagene), in which the sequence encoding SRAB may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors (Promega, Madison, WI) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews,

see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding SRAB may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L.E. in McGraw Hill *Yearbook of Science and Technology* (1992) McGraw Hill, New York, NY; pp. 191-196.

An insect system may also be used to express SRAB. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding SRAB may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of SRAB will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which SRAB may be expressed (Engelhard, E.K. et al. (1994) *Proc. Nat. Acad. Sci.* 91:3224-3227).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding SRAB may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing SRAB in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding SRAB. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding SRAB, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or

translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation
5 codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of
10 the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which
15 have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express SRAB may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression
20 elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using
25 tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1980) *Cell* 22:817-23) genes which can be employed in tk⁻ or aprt⁻ cells, respectively. Also, antimetabolite,
30 antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14) and als or pat, which confer resistance to chlorsulfuron and

phosphinotricin acetyltransferase, respectively (Murry, supra). Additional selectable genes have been described, for example, *trpB*, which allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-51). Recently, the use of visible markers has gained
5 popularity with such markers as anthocyanins, β glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. et al. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of
10 interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding SRAB is inserted within a marker gene sequence, recombinant cells containing sequences encoding SRAB can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding SRAB under the control of a single promoter. Expression of the marker gene in response to induction or selection
15 usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain the nucleic acid sequence encoding SRAB and express SRAB may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based
20 technologies for the detection and/or quantification of nucleic acid or protein.

The presence of polynucleotide sequences encoding SRAB can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or portions or fragments of polynucleotides encoding SRAB. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences encoding SRAB to detect transformants
25 containing DNA or RNA encoding SRAB. As used herein "oligonucleotides" or "oligomers" refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

A variety of protocols for detecting and measuring the expression of SRAB, using either
30 polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SRAB is preferred, but a

competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, MN) and Maddox, D.E. et al. (1983; J. Exp. Med. 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding SRAB include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding SRAB, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); and U.S. Biochemical Corp., Cleveland, OH). Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding SRAB may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode SRAB may be designed to contain signal sequences which direct secretion of SRAB through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding SRAB to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, CA) between the purification domain and SRAB may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing SRAB and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity

chromatography as described in Porath, J. et al. (1992, Prot. Exp. Purif. 3:263-281) while the enterokinase cleavage site provides a means for purifying SRAB from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D.J. et al. (1993; DNA Cell Biol. 12:441-453).

- 5 In addition to recombinant production, fragments of SRAB may be produced by direct peptide synthesis using solid-phase techniques Merrifield J. (1963) J. Am. Chem. Soc. 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of SRAB may be chemically synthesized
10 separately and combined using chemical methods to produce the full length molecule.

THERAPEUTICS

- Chemical and structural similarity exists among SRAB, rat Rab 26 protein, and human Rab 8. The expression of Rab proteins in pancreas, kidney, brain, salivary gland, and lung tissues and their association with cell cycle control and viral processing suggests that SRAB plays a role
15 in disorders associated with vesicle trafficking, viral infection, and cancer.

- Therefore, in one embodiment, SRAB or a fragment or derivative thereof may be administered to a subject to treat a disorder associated with vesicle trafficking. Such disorders may include, but are not limited to, autoimmune sialosis, choroideremia, cystic fibrosis, diabetes mellitus, hyperglycemia, hypoglycemia, tubulointerstitial nephritis, pancreatic enzyme deficiency,
20 pancreatitis, polycystic renal disease.

- In another embodiment, SRAB or a fragment or derivative thereof may be administered to a subject to treat a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma and sarcoma. Such cancers may include, but are not limited to, cancers of the endocrine, gastrointestinal and nervous systems such as insulinoma, gastrinoma, multiple endocrine neoplasia, and cancers of the
25 adrenal gland, brain, breast, colon, esophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland, prostate, salivary gland, stomach, thyroid, and uterus.

In another embodiment, a vector capable of expressing SRAB, or a fragment or a derivative thereof, may also be administered to a subject to treat a disorder associated with vesicle trafficking including those described above.

- 30 In another embodiment, a vector capable of expressing SRAB, or a fragment or a derivative thereof, may also be administered to a subject to treat a cancer including those described above.

In another embodiment, antagonists or inhibitors of SRAB may be administered to a

subject to prevent or treat a viral infection. Such viral infections may include, but are not limited to, those caused by retroviruses, particularly HIV and HTLV; hepadnaviruses, particularly hepatitis C; hantaviruses; herpesviruses, particularly HSV, EBV, and CMV; and arboviruses, particularly those causing encephalitis, yellow fever, and hemorrhagic fever.

5 In another embodiment, antagonists or inhibitors of SRAB may be administered to a subject to prevent or treat a disorder associated with vesicle trafficking. In the nervous system, uncontrolled or excessive vesicle trafficking may cause myoclonic phenomena such as shaking, fits, tremors, or palsy. Therefore, disorders of vesicle trafficking include, but not limited to, cerebral palsy, epilepsy, Huntington's disease, Parkinson's disease, petit and grand mal seizures,
10 and schizophrenia.

In one aspect, antibodies which are specific for SRAB may be used directly as an antagonist, or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express SRAB or in other situations where the decrease of SRAB levels is therapeutically desirable.

15 In another embodiment, a vector expressing the complementary or antisense sequence of the polynucleotide encoding SRAB may be administered to a subject to treat a viral infection including those described above.

In another embodiment, a vector expressing the complementary or antisense sequence of the polynucleotide encoding SRAB may be administered to a subject to treat a disorder associated
20 with vesicle trafficking including those described above.

In other embodiments, any of the therapeutic proteins, antagonists, antibodies, agonists, complementary or antisense sequences, or vectors described above may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to
25 conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Antagonists or inhibitors of SRAB may be produced using methods which are generally
30 known in the art. In particular, purified SRAB may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SRAB.

Antibodies specific for SRAB may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single

chain, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with SRAB or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the peptides, fragments, or oligopeptides used to induce antibodies to SRAB have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids. It is also preferable that they are identical to a portion of the amino acid sequence of the natural protein, and they may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of SRAB amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule.

Monoclonal antibodies to SRAB may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SRAB-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:11120-3).

Antibodies may also be produced by inducing in vivo production in the lymphocyte

population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for SRAB may also be
5 generated. For example, such fragments include, but are not limited to, the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 254:1275-1281).

10 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between SRAB and its specific antibody. A two-site, monoclonal-based immunoassay utilizing
15 monoclonal antibodies reactive to two non-interfering SRAB epitopes is preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides encoding SRAB, or any fragment thereof, or antisense molecules, may be used for therapeutic purposes. In one aspect, antisense to the polynucleotide encoding SRAB may be used in situations in which it would be
20 desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding SRAB. Thus, antisense molecules may be used to modulate SRAB activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding
25 SRAB.

Expression vectors derived from retro viruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express antisense molecules complementary to
30 the polynucleotides of the gene encoding SRAB. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra).

Genes encoding SRAB can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide or fragment thereof which encodes SRAB.

Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements
5 are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA, or PNA, to the control regions of the gene encoding SRAB, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly,
10 inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt.
15 Kisco, NY). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage.
20 Examples which may be used include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding SRAB.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20
25 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Antisense molecules and ribozymes of the invention may be prepared by any method
30 known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding SRAB. Such DNA sequences may be incorporated into a wide

variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible
5 modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified
10 forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into
15 that same patient. Delivery by transfection and by liposome injections may be achieved using methods which are well known in the art.

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

20 An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of SRAB, antibodies to SRAB, mimetics, agonists, antagonists, or inhibitors of SRAB. The compositions may be administered alone or in combination with at least one other agent, such as
25 stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any
30 number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain

suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, 5 PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by 10 the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, 15 mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

20 Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

25 Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty 30 oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain

substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

10 The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of SRAB, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example SRAB or fragments thereof, antibodies of SRAB, agonists, antagonists or inhibitors of SRAB, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be

determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind SRAB may be used for the diagnosis of conditions or diseases characterized by expression of SRAB, or in assays to monitor patients being treated with SRAB, agonists, antagonists or inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for SRAB include methods which utilize the antibody and a label to detect SRAB in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used, several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring SRAB are known in the art and provide a basis for diagnosing altered or abnormal levels of SRAB expression. Normal or standard values for SRAB expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to SRAB
5 under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of SRAB expressed in subject, control and disease, samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

10 In another embodiment of the invention, the polynucleotides encoding SRAB may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of SRAB may be correlated with disease. The diagnostic assay may be used to distinguish between absence,
15 presence, and excess expression of SRAB, and to monitor regulation of SRAB levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding SRAB or closely related molecules, may be used to identify nucleic acid sequences which encode SRAB. The specificity
20 of the probe, whether it is made from a highly specific region, e.g., 10 unique nucleotides in the 5' regulatory region, or a less specific region, e.g., especially in the 3' coding region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding SRAB, alleles, or related sequences.

25 Probes may also be used for the detection of related sequences, and should preferably contain at least 50% of the nucleotides from any of the SRAB encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence of SEQ ID NO:2 or from genomic sequence including promoter, enhancer elements, and introns of the naturally occurring SRAB.

30 Means for producing specific hybridization probes for DNAs encoding SRAB include the cloning of nucleic acid sequences encoding SRAB or SRAB derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA

polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as ^{32}P or ^{35}S , or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

- 5 Polynucleotide sequences encoding SRAB may be used for the diagnosis of conditions or diseases which are associated with expression of SRAB. Examples of such conditions or diseases include disorders associated with defects in the secretory and digestive systems. Such conditions and diseases may include, but are not limited to, autoimmune sialosis, choroideremia, cystic fibrosis, diabetes mellitus, hyperglycemia, hypoglycemia, tubulointerstitial nephritis,
- 10 pancreatic enzyme deficiency, pancreatitis, polycystic renal disease, insulinoma, gastrinoma, multiple endocrine neoplasia, adenocarcinoma, leukemia, lymphoma, melanoma, sarcoma, and cancers of the adrenal gland, brain, breast, colon, esophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland, prostate, salivary gland, stomach, thyroid, and uterus.

- In another embodiment, antagonists or inhibitors of SRAB may be administered to a
- 15 subject to prevent or treat a disorder associated with vesicle trafficking. In the nervous system, excessive signal transduction due to increased vesicle trafficking may include, but is not limited to, cerebral palsy, epilepsy, Huntington's disease, Parkinson's disease, petit and grand mal seizures, and schizophrenia.

- In one aspect, antibodies which are specific for SRAB may be used directly as an
- 20 antagonist, or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express SRAB or in other situations where the decrease of SRAB levels is therapeutically desirable.

- Polynucleotide sequences encoding SRAB may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA
- 25 or chip assays utilizing fluids or tissues from patient biopsies to detect altered SRAB expression. Such qualitative or quantitative methods are well known in the art.

- In a particular aspect, the nucleotide sequences encoding SRAB may be useful in assays that detect activation or induction of various cancers, particularly those mentioned above. The nucleotide sequences encoding SRAB may be labeled by standard methods, and added to a fluid
- 30 or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable control sample, the nucleotide sequences have

hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding SRAB in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of disease associated with expression of SRAB, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, which encodes SRAB, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease.

Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding SRAB may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5'→3') and another with antisense (3'←5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of SRAB include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated (Melby, P.C. et al. (1993) J. Immunol. Methods, 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236). The speed
5 of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In another embodiment of the invention, the nucleic acid sequences which encode SRAB may also be used to generate hybridization probes which are useful for mapping the naturally
10 occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS, or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C.M. (1993) Blood Rev. 7:127-134, and Trask, B.J. (1991) Trends Genet.
15 7:149-154.

FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, NY) may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene
20 encoding SRAB on a physical chromosomal map and a specific disease, or predisposition to a specific disease, may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier, or affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such
25 as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for
30 disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R.A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide

sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.

In another embodiment of the invention, SRAB, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, 5 affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between SRAB and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in 10 published PCT application WO84/03564. In this method, as applied to SRAB large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with SRAB, or fragments thereof, and washed. Bound SRAB is then detected by methods well known in the art. Purified SRAB can also be coated directly onto plates for use in the aforementioned drug screening techniques.

15 Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding SRAB specifically compete with a test compound for binding SRAB. In this manner, the antibodies can be used to detect the presence of any peptide 20 which shares one or more antigenic determinants with SRAB.

In additional embodiments, the nucleotide sequences which encode SRAB may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

25 The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

INDUSTRIAL APPLICABILITY

I PANCNOT04 cDNA Library Construction

The PANCNOT04 cDNA library was constructed from pancreatic tissue from a five year 30 old boy who had been killed in a motor vehicle accident (specimen #GG359; International Institute for Advanced Medicine, Exton PA).

The frozen tissue was homogenized and lysed using a Brinkmann Homogenizer Polytron PT-3000 (Brinkmann Instruments, Westbury NJ) in guanidinium isothiocyanate solution. The

lysate was centrifuged over a 5.7 M CsCl cushion using an Beckman SW28 rotor in a Beckman L8-70M Ultracentrifuge (Beckman Instruments) for 18 hours at 25.000 rpm at ambient temperature. The RNA was extracted once with phenol chloroform pH 8.0, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in RNase-free water and DNase treated at 37°C. Extraction and precipitation were repeated as before. The mRNA was isolated with the Qiagen Oligotex kit (QIAGEN Inc, Chatsworth CA) and used to construct the cDNA library.

The mRNA was handled according to the recommended protocols in the SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning (Catalog #18248-013, Gibco BRL, Gaithersburg, MD). cDNAs were fractionated on a Sepharose CL4B column (Catalog #275105-01, Pharmacia), and those cDNAs exceeding 400 bp were ligated into pSport I. The plasmid pSport I was subsequently transformed into DH5a™ competent cells (Catalog #18258-012, Gibco BRL).

II Isolation and Sequencing of cDNA Clones

Plasmid DNA was released from the cells and purified using the REAL Prep 96 Plasmid Kit (Catalog #26173; QIAGEN, Inc). This kit enables the simultaneous purification of 96 samples in a 96-well block using multi-channel reagent dispensers. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Catalog #22711, GIBCO/BRL) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) the cultures were incubated for 19 hours after the wells were inoculated and then lysed with 0.3 ml of lysis buffer; 3) following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a Beckman 96-well block for storage.

The cDNAs were sequenced by the method of Sanger F and AR Coulson (1975; J Mol Biol 94:441f), using a Hamilton Micro Lab 2200 (Hamilton, Reno NV) in combination with Peltier Thermal Cyclers (PTC200 from MJ Research, Watertown MA) and Applied Biosystems 377 or 373 DNA Sequencing Systems; and the reading frame was determined.

III Similarity Searching of cDNA Clones and Their Deduced Proteins

The nucleotide sequences of the Sequence Listing or amino acid sequences deduced from them were used as query sequences against databases such as GenBank, SwissProt, BLOCKS, and Pima II. These databases which contain previously identified and annotated sequences were searched for regions of similarity using BLAST, which stands for Basic Local Alignment Search Tool (Altschul (1993) supra, Altschul (1990) supra).

BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal or plant) origin. Other algorithms such as the one described in Smith
5 RF and TF Smith (1992; Protein Engineering 5:35-51), incorporated herein by reference, can be used when dealing with primary sequence patterns and secondary structure gap penalties. As disclosed in this application, the sequences have lengths of at least 49 nucleotides, and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach, as detailed in Karlin and Altschul (supra) and incorporated herein
10 by reference, searches for matches between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. In this application, threshold was set at 10^{-25} for nucleotides and 10^{-14} for peptides.

Incyte nucleotide sequences were searched against the GenBank databases for primate
15 (pri), rodent (rod), and mammalian sequences (mam), and deduced amino acid sequences from the same clones are searched against GenBank functional protein databases, mammalian (mamp), vertebrate (vrtp) and eukaryote (eukp), for similarity. The relevant database for a particular match were reported as a G1xxx±p (where xxx is pri, rod, etc and if present, p = peptide). Product score, the calculation of which is shown below, was used to determine the electronic
20 stringency. For an exact match, product score was set at 70 with a conservative lower limit set at approximately 40 (1-2% error due to uncalled bases).

IV Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which
25 RNAs from a particular cell type or tissue have been bound (Sambrook et al., supra).

Analogous computer techniques using BLAST (Altschul, S.F. 1993 and 1990, supra) are used to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ™ database (Incyte Pharmaceuticals). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be
30 modified to determine whether any particular match is categorized as exact or homologous.

The basis of the search is the product score which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

100

-35-

SUBSTITUTE SHEET (RULE 26)

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower 5 scores may identify related molecules.

The results of northern analysis are reported as a list of libraries in which the transcript encoding SRAB occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the cDNA 10 library.

V Extension of SRAB-Encoding Polynucleotides

Nucleic acid sequence of Incyte Clone 738957 was used to design oligonucleotide primers for extending the nucleotide sequence to full length. One primer was synthesized to initiate extension in the antisense direction, and the other was synthesized to extend sequence in the 15 sense direction. Primers were used to facilitate the extension of the known sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest. The initial primers were designed from the cDNA using OLIGO 4.06 (National Biosciences), or another appropriate program, to be about 22 to about 30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures of about 68° to 20 about 72° C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries (Gibco/BRL) were used to extend the sequence. If more than one extension is necessary or desired, additional sets of primers are designed to further extend the known region.

25 High fidelity amplification was obtained by following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix. Beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, PCR was performed using the Peltier Thermal Cycler (PTC200; M.J. Research, Watertown, MA) and the following parameters:

30	Step 1	94° C for 1 min (initial denaturation)
	Step 2	65° C for 1 min
	Step 3	68° C for 6 min
	Step 4	94° C for 15 sec
	Step 5	65° C for 1 min

- | | | |
|---|---------|--|
| | Step 6 | 68° C for 7 min |
| | Step 7 | Repeat step 4-6 for 15 additional cycles |
| | Step 8 | 94° C for 15 sec |
| | Step 9 | 65° C for 1 min |
| 5 | Step 10 | 68° C for 7:15 min |
| | Step 11 | Repeat step 8-10 for 12 cycles |
| | Step 12 | 72° C for 8 min |
| | Step 13 | 4° C (and holding) |

10 A 5-10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were excised from the gel, purified using QIAQuick™ (QIAGEN Inc., Chatsworth, CA), and trimmed of overhangs using Klenow enzyme to facilitate religation and cloning.

15 After ethanol precipitation, the products were redissolved in 13 μ l of ligation buffer, 1 μ l T4-DNA ligase (15 units) and 1 μ l T4 polynucleotide kinase were added, and the mixture was incubated at room temperature for 2-3 hours or overnight at 16° C. Competent *E. coli* cells (in 40 μ l of appropriate media) were transformed with 3 μ l of ligation mixture and cultured in 80 μ l of SOC medium (Sambrook et al., supra). After incubation for one hour at 37° C, the *E. coli*
 20 mixture was plated on Luria Bertani (LB)-agar (Sambrook et al., supra) containing 2x Carb. The following day, several colonies were randomly picked from each plate and cultured in 150 μ l of liquid LB/2x Carb medium placed in an individual well of an appropriate, commercially-available, sterile 96-well microtiter plate. The following day, 5 μ l of each overnight culture was transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5 μ l of each sample
 25 was transferred into a PCR array.

For PCR amplification, 18 μ l of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction were added to each well. Amplification was performed using the following conditions:

- | | | |
|----|--------|--|
| 30 | Step 1 | 94° C for 60 sec |
| | Step 2 | 94° C for 20 sec |
| | Step 3 | 55° C for 30 sec |
| | Step 4 | 72° C for 90 sec |
| | Step 5 | Repeat steps 2-4 for an additional 29 cycles |
| 35 | Step 6 | 72° C for 180 sec |
| | Step 7 | 4° C (and holding) |

Aliquots of the PCR reactions were run on agarose gels together with molecular weight markers. The sizes of the PCR products were compared to the original partial cDNAs, and

appropriate clones were selected, ligated into plasmid, and sequenced.

VI Labeling and Use of Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham) and T4 polynucleotide kinase (DuPont NEN[®], Boston, MA). The labeled oligonucleotides are substantially purified with Sephadex G-25 superfine resin column (Pharmacia & Upjohn). A portion containing 10^7 counts per minute of each of the sense and antisense oligonucleotides is used in a typical membrane based hybridization analysis of human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II; DuPont NEN[®]).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham, NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR[™] film (Kodak, Rochester, NY) is exposed to the blots in a Phosphorimager cassette (Molecular Dynamics, Sunnyvale, CA) for several hours, hybridization patterns are compared visually.

VII Antisense Molecules

Antisense molecules or sequence complementary to the SRAB-encoding sequence, or any part thereof, is used to inhibit in vivo or in vitro expression of naturally occurring SRAB. Although use of antisense oligonucleotides, comprising about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. An oligonucleotide based on the coding sequences of SRAB, as shown in Figures 1A, 1B, 1C and 1D, is used to inhibit expression of naturally occurring SRAB. The complementary oligonucleotide is designed from the most unique 5' sequence as shown in Figures 1A, 1B, 1C and 1D and used either to inhibit transcription by preventing promoter binding to the upstream nontranslated sequence or translation of an SRAB-encoding transcript by preventing the ribosome from binding. Using an appropriate portion of the signal and 5' sequence of SEQ ID NO:2, an effective antisense oligonucleotide includes any 15-20 nucleotides spanning the region which translates into the signal or 5' coding sequence of the polypeptide as shown in Figures 1A,

1B, 1C and 1D.

VIII Expression of SRAB

Expression of SRAB is accomplished by subcloning the cDNAs into appropriate vectors and transforming the vectors into host cells. In this case, the cloning vector, pSport, previously
5 used for the generation of the cDNA library is used to express SRAB in *E. coli*. Upstream of the cloning site, this vector contains a promoter for β -galactosidase, followed by sequence containing the amino-terminal Met, and the subsequent seven residues of β -galactosidase. Immediately following these eight residues is a bacteriophage promoter useful for transcription and a linker containing a number of unique restriction sites.

10 Induction of an isolated, transformed bacterial strain with IPTG using standard methods produces a fusion protein which consists of the first eight residues of β -galactosidase, about 5 to 15 residues of linker, and the full length protein. The signal residues direct the secretion of SRAB into the bacterial growth media which can be used directly in the following assay for activity.

15 IX Demonstration of SRAB Activity

SRAB can be expressed in a mammalian cell line such as 293T by transfecting with an eukaryotic expression vector encoding SRAB. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A small amount of a second plasmid, which expresses any one of a number of reporter genes
20 such as β -galactosidase, is co-transformed into the cells in order to allow rapid identification of those cells which have taken up and expressed the foreign DNA. The cells are cultured in a defined synthetic medium with concentrations of GTP for at least 48 hours after transformation to allow expression and accumulation of SRAB and β -galactosidase.

Transformed cells expressing β -galactosidase are stained blue when a suitable
25 colorimetric substrate is added to the culture media under conditions that are well known in the art. Increasing concentrations of GTP induces increasing numbers of reporter gene positive cells (Ren, M. et al. (1996) Proc. Natl. Acad. Sci. 93:5151-5155). GTP-treated cells which were not transformed with the SRAB expression vector are used as controls as are SRAB transfected cells cultured without supplemental GTP.

30 X Production of SRAB Specific Antibodies

SRAB that is substantially purified using PAGE electrophoresis (Sambrook, supra), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence deduced from SEQ ID NO:2 is analyzed using

DNASTAR software (DNASTAR Inc) to determine regions of high immunogenicity and a corresponding oligopolypeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions, is described by Ausubel et al. (supra), and others.

- 5 Typically, the oligopeptides are 15 residues in length, synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma, St. Louis, MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel et al., supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for
10 antipeptide activity, for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

XI Purification of Naturally Occurring SRAB Using Specific Antibodies

- Naturally occurring or recombinant SRAB is substantially purified by immunoaffinity chromatography using antibodies specific for SRAB. An immunoaffinity column is constructed
15 by covalently coupling SRAB antibody to an activated chromatographic resin, such as CnBr-activated Sepharose (Pharmacia & Upjohn). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

- Media containing SRAB is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SRAB (e.g., high ionic
20 strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SRAB binding (eg, a buffer of pH 2-3 or a high concentration of a chaotrope, such as urea or thiocyanate ion), and SRAB is collected.

XII Identification of Molecules Which Interact with SRAB

- SRAB or biologically active fragments thereof are labeled with ¹²⁵I Bolton-Hunter
25 reagent (Bolton et al. (1973) Biochem. J. 133:529). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled SRAB, washed and any wells with labeled SRAB complex are assayed. Data obtained using different concentrations of SRAB are used to calculate values for the number, affinity, and association of SRAB with the candidate molecules.

- 30 All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred

embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: INCYTE PHARMACEUTICALS, INC.
- (ii) TITLE OF THE INVENTION: NOVEL RAB PROTEIN
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 - (B) STREET: 3174 Porter Drive
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 - (A) PCT APPLICATION NUMBER: To Be Assigned
 - (B) FILING DATE: Herewith
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/824,873
 - (B) FILING DATE: 26-MAR-1997
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Billings, Lucy J.
 - (B) REGISTRATION NUMBER: 36,749
 - (C) REFERENCE/DOCKET NUMBER: PF-0240 PCT
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 650-855-0555
 - (B) TELEFAX: 650-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 190 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: PANCNOT04
- (B) CLONE: 738957

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Leu	Val	Gly	Asp	Ser	Gly	Val	Gly	Lys	Thr	Cys	Leu	Leu	Gly	Ala
1				5					10					15	
Ile	Gln	Gly	Trp	Cys	Phe	Pro	Gly	Gly	Asp	Leu	His	Leu	His	Arg	Ser
			20					25					30		
Ile	Asp	Phe	Arg	Asn	Lys	Val	Leu	Asp	Val	Asp	Gly	Val	Lys	Val	Lys
		35				40					45				
Leu	Gln	Met	Trp	Asp	Thr	Ala	Gly	Gln	Glu	Arg	Phe	Arg	Ser	Val	Thr
		50			55					60					
His	Ala	Tyr	Tyr	Arg	Asp	Ala	His	Ala	Leu	Leu	Leu	Leu	Tyr	Asp	Val
65				70					75				80		
Thr	Asn	Lys	Ala	Ser	Phe	Asp	Asn	Ile	Gln	Ala	Trp	Leu	Thr	Glu	Ile
			85					90					95		

```

His Glu Tyr Ala Gln His Asp Val Ala Leu Met Leu Leu Gly Asn Lys
      100      105      110
Val Asp Ser Ala His Glu Arg Val Val Lys Arg Glu Asp Gly Glu Lys
      115      120      125
Leu Ala Lys Glu Tyr Gly Leu Pro Phe Met Glu Thr Ser Ala Lys Thr
      130      135      140
Gly Leu Asn Val Asp Leu Ala Phe Thr Ala Ile Ala Lys Glu Leu Lys
      145      150      155      160
Gln Arg Ser Met Lys Ala Pro Ser Glu Pro Arg Phe Arg Leu His Asp
      165      170      175
Tyr Val Lys Arg Glu Gly Arg Gly Ala Ser Cys Cys Arg Pro
      180      185      190

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1340 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: PANCNOT04
 (B) CLONE: 738957

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

CTTGGCGGGCG GTGTCGACTT CTACGACGTC GCCTTCAAGG TCATGCTGGT GGGGGACTCG      60
GGTGTGGGGA AGACCTGTCT GCTGGGTGCG ATTCAAGGAT GGTGCTTTCC TGGCGGGGAC      120
CTTCATCTCC ACCGTAGCAT TGACTTCCGG AACAAAGTTC TGGACGTGGA TGGTGTGAAG      180
GTGAAGCTGC AGATGTGGGA CACAGCTGGT CAGGAGCGGT TCCGCAGTGT TACCCATGCC      240
TACTACCGGG ATGCTCATGC TCTGCTGCTG CTCTACGATG TCACCAACAA GGCCTCCTTT      300
GACAACATCC AGGCCTGGCT GACCGAGATC CACGAGTACG CCCAGCACGA CGTGGCGCTC      360
ATGCTGCTGG GGAACAAGGT GGACTCTGCC CATGAGCGTG TGGTGAAGAG GGAGGACGGG      420
GAGAAGCTGG CCAAGGAGTA TGGACTGCCC TTCATGGAGA CCAGCGCCAA GACGGGCCTC      480
AACGTGGACT TGGCCTTCAC AGCCATAGCA AAGGAGTTGA AGCAGCGCTC CATGAAGGCT      540
CCCAGCGAGC CGCGCTTCCG GCTGCATGAT TACGTTAAGA GGGAGGGTCG AGGGGCCTCC      600
TGCTGCCGCC CTTGAACCTG GCTGAGCTCA GTCTCTGGA GGAGGCCGCC CAGTCCCTAG      660
AAGGCTGGAC AGAGGGTCTC CAGGCCCTTC TGACTTTGTT GCCCAGTGGC CAACGCCCGA      720
GTGTCTGTTT TCAGGAGCCC CAGGTCAAGC CTTGTCCCTT CCTCCTCCCA GCAACAGTCC      780
CAACAAGCAG GCTTCTGAGA GCCCGTGGCC GCACACTGGC CGCCACGGAA AAGCAGTCTT      840
CTGCACGGGA CGGGGAGCGG CAAGTGGACA GACTTTGCCA CGGTGCTCTG CTGCCCCCTC      900
CTGGGCACGT CCAGGTGAGG GAGGGCTGGG GCTGGCACCA CGCACAGTGC CTAACCCTAG      960
AAAAGCCATG TCTTCAGCCG CACATGCCCC GGCAGCTAAG GGAGGACGCC TGCCACGCC      1020
TGGGACAGAA GGCTTCACTG CTAATCACAT CGTGCATCTG TGTGTCCTGG GAGCTGCCTG      1080
CTCCCGGCC ACCCTCTAGG AGGCTCTGGC TCAAACAGCA ATAGGGTCTT CCTCACTGAC      1140
CTTGGAGGAT GCCTGTGGCC TTGTGATAAA ATGTGGGAAA TCACAGAAAA CACCAGAAAC      1200
AACAAC TGCC AGCCCGGCCT GGCCACAGGT GAGGTCTGTG ATTTCCGAGC ACGCTCCACC      1260
TTGCACTCAA CTTGGCCTTT TGATTGCACA AGCCTTTGTT TTCAGTCCTA GTGAATAAAG      1320
TTGTGTTTTT TGGAAAAAAA

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 190 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
 (B) CLONE: 619734

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Leu Val Gly Asp Ser Gly Val Gly Lys Thr Cys Leu Leu Val Arg
  1           5           10           15

```

```

Phe Lys Asp Gly Ala Phe Leu Ala Gly Thr Phe Ile Ser Thr Val Gly
  20                      25                      30
Ile Asp Phe Arg Asn Lys Val Leu Asp Val Asp Gly Met Lys Val Lys
  35                      40                      45
Leu Gln Ile Trp Asp Thr Ala Gly Gln Glu Arg Phe Arg Ser Val Thr
  50                      55                      60
His Ala Tyr Tyr Arg Asp Ala His Ala Leu Leu Leu Leu Tyr Asp Ile
  65                      70                      75                      80
Thr Asn Lys Asp Ser Phe Asp Asn Ile Gln Ala Trp Leu Thr Glu Ile
  85                      90                      95
Gln Glu Tyr Ala Gln Gln Asp Val Val Leu Met Leu Leu Gly Asn Lys
  100                     105                     110
Val Asp Ser Thr Gln Glu Arg Val Val Lys Arg Glu Asp Gly Glu Lys
  115                     120                     125
Leu Ala Lys Glu Tyr Gly Leu Pro Phe Met Glu Thr Ser Ala Lys Ser
  130                     135                     140
Gly Leu Asn Val Asp Leu Ala Phe Thr Ala Ile Ala Lys Glu Leu Lys
  145                     150                     155                     160
Gln Arg Ser Thr Lys Ala Pro Ser Glu Pro Arg Phe Arg Leu His Asp
  165                     170                     175
Tyr Val Lys Arg Glu Gly Arg Gly Val Ser Cys Cys Arg Leu
  180                     185                     190

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 234746

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Ala Lys Thr Tyr Asp Tyr Leu Phe Lys Leu Leu Leu Ile Gly Asp
  1                      5                      10                      15
Ser Gly Val Gly Lys Thr Cys Val Leu Phe Arg Phe Ser Glu Asp Ala
  20                      25                      30
Phe Asn Ser Thr Phe Ile Ser Thr Ile Gly Ile Asp Phe Lys Ile Arg
  35                      40                      45
Thr Ile Glu Leu Asp Gly Lys Arg Ile Lys Leu Gln Ile Trp Asp Thr
  50                      55                      60
Ala Gly Gln Glu Arg Phe Arg Thr Ile Thr Thr Ala Tyr Tyr Arg Gly
  65                      70                      75                      80
Ala Met Gly Ile Met Leu Val Tyr Asp Ile Thr Asn Glu Lys Ser Phe
  85                      90                      95
Asp Asn Ile Arg Asn Trp Ile Arg Asn Ile Glu Glu His Ala Ser Ala
  100                     105                     110
Asp Val Glu Lys Met Ile Leu Gly Asn Lys Cys Asp Val Asn Asp Lys
  115                     120                     125
Arg Gln Val Ser Lys Glu Arg Gly Glu Lys Leu Ala Leu Asp Tyr Gly
  130                     135                     140
Ile Lys Phe Met Glu Thr Ser Ala Lys Ala Asn Ile Asn Val Glu Asn
  145                     150                     155                     160
Ala Phe Phe Thr Leu Ala Arg Asp Ile Lys Ala Lys Met Asp Lys Lys
  165                     170                     175
Leu Glu Gly Asn Ser Pro Gln Gly Ser Asn Gln Gly Val Lys Ile Thr
  180                     185                     190
Pro Asp Gln Lys Arg Ser Ser Phe Phe Arg Cys Val Leu
  195                     200                     205

```

What is claimed is:

1. A substantially purified Rab protein comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof.
2. An isolated and purified polynucleotide sequence encoding the Rab protein of claim 1.
3. A polynucleotide sequence which hybridizes under stringent conditions to the polynucleotide sequence of claim 2.
4. A hybridization probe comprising the polynucleotide sequence of claim 2.
5. An isolated and purified polynucleotide sequence comprising SEQ ID NO:2 or variants thereof.
6. A polynucleotide sequence which is complementary to the polynucleotide sequence of claim 2 or variants thereof.
7. A hybridization probe comprising the polynucleotide sequence of claim 6.
8. An expression vector containing the polynucleotide sequence of claim 2.
9. A host cell containing the vector of claim 8.
10. A method for producing a polypeptide comprising the amino acid sequence of SEQ ID NO:1 the method comprising the steps of:
 - a) culturing the host cell of claim 9 under conditions suitable for the expression of the polypeptide; and
 - b) recovering the polypeptide from the host cell culture.
11. A pharmaceutical composition comprising a substantially purified RAB protein having an amino acid sequence of claim 1 in conjunction with a suitable pharmaceutical carrier.
12. A purified antibody which binds specifically to the polypeptide of claim 1.
13. A purified agonist which specifically binds to and modulates the activity of the polypeptide of claim 1.
14. A purified antagonist which specifically binds to and modulates the activity of the polypeptide of claim 1.
15. A method for treating a disorder associated with vesicle trafficking comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 11.
16. A method for treating cancer comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 11.

17. A method for treating a viral infection, comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 14.

18. A method for treating a disorder associated with vesicle trafficking, comprising administering to a subject in need of such treatment an effective amount of the antagonist of
5 claim 14.

19. A method for detection of a polynucleotide which encodes a RAB protein in a biological sample comprising the steps of:

- a) hybridizing the polynucleotide of claim 6 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and
- 10 b) detecting said hybridization complex, wherein the presence of said complex correlates with the presence of a polynucleotide encoding a protein in said biological sample.

CTT	GGC	GGC	GGT	GTC	GAC	TTC	TAC	GAC	GTC	GCC	TTC	AAG	GTC	ATG	CTG	GTG	GGG
9		18			27					36				45			54
														M	L	V	G
GAC	TCG	GGT	GTG	GGG	AAG	ACC	TGT	CTG	CTG	GGT	GCG	ATT	CAA	GGA	TGG	TGC	TTT
63		72			81					90				99			108
D	S	G	V	G	K	T	C	L	L	G	A	I	Q	G	W	C	F
CCT	GGC	GGG	GAC	CTT	CAT	CTC	CAC	CGT	AGC	ATT	GAC	TTC	CGG	AAC	AAA	GTT	CTG
117		126			135					144				153			162
P	G	G	D	L	H	L	H	R	S	I	D	F	R	N	K	V	L
GAC	GTG	GAT	GGT	GTG	AAG	GTG	AAG	CTG	CAG	ATG	TGG	GAC	ACA	GCT	GGT	CAG	GAG
171		180			189					198				207			216
D	V	G	V	K	L	V	K	L	Q	M	W	D	T	A	G	Q	E
CGG	TTC	CGC	AGT	GTT	ACC	CAT	GCC	TAC	TAC	CGG	GAT	GCT	CAT	GCT	CTG	CTG	CTG
225		234			243					252				261			270
R	F	R	S	V	T	H	A	Y	Y	R	D	A	H	A	L	L	L
CTC	TAC	GAT	GTC	ACC	AAC	AAG	GCC	TCC	TTT	GAC	AAC	ATC	CAG	GCC	TGG	CTG	ACC
279		288			297					306				315			324
L	Y	D	V	T	N	K	A	S	F	D	N	I	Q	A	W	L	T
GAG	ATC	CAC	GAG	TAC	GCC	CAG	CAC	GAC	GTG	GCG	CTC	ATG	CTG	CTG	GGG	AAC	AAG
333		342			351					360				369			378
E	I	H	E	Y	A	Q	H	D	V	A	L	M	L	L	G	N	K

FIGURE 1A

387	GTG GAC TCT GCC CAT GAG CGT GTG GTG AAG AGG GAG GAC GGG GAG AAG CTG GCC	414	423	432
	V D S A H E R V V K R E D G E K L A			
441	AAG GAG TAT GGA CTG CCC TTC ATG GAG ACC AGC GCC AAG ACG GGC CTC AAC GTG	468	477	486
	K E Y G L P F M E T S A K T G L N V			
495	GAC TTG GCC TTC ACA GCC ATA GCA AAG GAG TTG AAG CAG CGC TCC ATG AAG GCT	513	522	531
	D L A F T A I A K E L K Q R S M K A			
549	CCC AGC GAG CCG CGC TTC CGG CTG CAT GAT TAC GTT AAG AGG GAG GGT CGA GGG	567	576	585
	P S E P R F R L H D Y V K R E G R G			
603	GCC TCC TGC TGC CGC CCT TGA ACC TGG CTG AGC TCA GTC CTC TGG AGG AGG CCG	621	630	639
	A S C C R P *			
657	CCC AGT CCC TAG AAG GCT GGA CAG AGG GTC TCC AGG CCC TTC TGA CTT TGT TGC	666	675	684
711	CCA GTG GCC AAC GCC CGA GTG TCT GTT TTC AGG AGC CCC AGG TCA AGC CTT GTG	720	729	738
765	CCT TCC TCC TCC CAG CAA CAG TCC CAA CAA GCA GGC TTC TGA GAG CCC GTG GCC	774	783	792
			801	810

FIGURE 1B

819	828	837	846	855	864
GCA CAC TGG CCG CCA CGG AAA AGC AGT CTT CTG CAC GGG ACG GGC AGC GGC AAG					
873	882	891	900	909	918
TGG ACA GAC TTT GCC ACG GTG CTC TGC TGC CCC CTC CTG GGC ACG TCC AGG TGA					
927	936	945	954	963	972
GGG AGG GCT GGG GCT GGC ACC ACG CAC AGT GCC TAA CCC TAG AAA AGC CAT GTC					
981	990	999	1008	1017	1026
TTC AGC CGC ACA TGC CCA GGC AGC TAA GGG AGG ACG CCT GCC CAC GCC TGG GAC					
1035	1044	1053	1062	1071	1080
AGA AGG CTT CAC TGC TAA TCA CAT CGT GCA TCT GTG TGT CCT GGC AGC TGC CTG					
1089	1098	1107	1116	1125	1134
CTC CCG GCC CAC CCT CTA GGA GGC TCT GGC TCA AAC AGC AAT AGG GTC TTC CTC					
1143	1152	1161	1170	1179	1188
ACT GAC CTT GGA GGA TGC CTG CTG TGG CCT TGT GAT AAA ATG TGG GAA ATC ACA GAA					
1197	1206	1215	1224	1233	1242
AAC ACC AGA AAC AAC AAC TGC CAG CCC GGC CTG GCC ACA GGT GAG GTC TGT GAT					
1251	1260	1269	1278	1287	1296
TTC CGA GCA CGC TCC ACC TTG CAC TCA ACT TGG CCT TTT GAT TGC ACA AGC CTT					

FIGURE 1C

1305 TGT TTT CAG TCC TAG TGA ATA AAG TTG TGT TTT CTG GAA AAA AA
1314
1323
1332

FIGURE 1D

140	T	S	A	K	T	G	L	N	V	D	L	A	F	T	A	I	A	K	E	L	K	Q	R	S	M	K	-	-	-	-	738957
140	T	S	A	K	S	G	L	N	V	D	L	A	F	T	A	I	A	K	E	L	K	Q	R	S	T	K	-	-	-	-	GI 619734
150	T	S	A	K	A	N	I	N	V	E	N	A	F	T	L	A	R	D	I	K	A	K	M	D	K	K	L	E	G	GI 234746	
166	-	A	P	S	E	P	R	F	R	L	H	D	Y	V	K	R	E	G	R	G	A	-	-	S	C	C	R	P	738957		
166	-	A	P	S	E	P	R	F	R	L	H	D	Y	V	K	R	E	G	R	G	V	-	-	S	C	C	R	L	GI 619734		
180	N	S	P	Q	G	S	N	Q	G	V	K	I	T	P	D	Q	Q	K	R	S	S	F	F	R	C	V	L	L	GI 234746		

FIGURE 2B

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/05871

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C07K14/47 C07K16/18 C12Q1/68 C12N15/11
A61K38/17 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K C12Q A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	WAGNER A. C. C. ET AL.: "Molecular cloning of a new member of the Rab protein family, Rab26, from rat pancreas." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 207, no. 3, 1995, pages 950-956, XP002072106	1-15, 18, 19
Y	see the whole document	17
Y	BOGERD H. P. ET AL.: "IDENTIFICATION OF A NOVEL CELLULAR COFACTOR FOR THE REV/REX CLASS OF RETROVIRAL REGULATORY PROTEINS" CELL, vol. 82, 11 August 1995, pages 485-494, XP002021069 see the whole document	17
-/--		

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

20 July 1998

Date of mailing of the international search report

31 jul 1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Mandl, B

INTERNATIONAL SEARCH REPORT

In International Application No

PCT/US 98/05871

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
X	<p>EMBL Database entry T08521; Accession number T08521; 6. August 1993; Adams et al.: 'Rapid cDNA sequencing from a directionally cloned human infant brain cDNA library.' XP002072108 see abstract</p> <p style="text-align: center;">---</p>	2-7
X	<p>EMBL Database entry HS132292; Accession number N48132; 18. February 1996; Hillier et al.: 'The WashU-Merck EST project.' XP002072117 see abstract</p> <p style="text-align: center;">---</p>	2-7
X	<p>EMBL Database entry HSR88060; Accession number R88060; 10. October 1995; Hillier et al.: 'The WashU-Merck EST project.' XP002072109 see abstract</p> <p style="text-align: center;">---</p>	3-7
A	<p>NOVICK P. AND BRENNWALD P.: "Friends and Family: The role of Rab GTPases in vesicular traffic." CELL, vol. 75, 1993, pages 597-601, XP002072107 cited in the application see the whole document</p> <p style="text-align: center;">-----</p>	1-19